

Data Mining Penjualan Produk dengan Algoritma Apriori Pada Jims Honey Pangkalpinang

Parlia Romadiana^{1*}, Delphia Wahyuningsih², Cindy Aprilia³, Amanda Monica⁴,
Anisah⁵, Syafrul Irawadi⁶

^{1,2,3,4,5,6}Institut Sains dan Bisnis Atma Luhur, Indonesia

E-mail: ¹parliaromadiana@atmaluhur.ac.id, ²delphibabel@atmaluhur.ac.id,

³1922500162@mahasiswa.atmaluhur.ac.id,

⁴1922500022@mahasiswa.atmaluhur.ac.id, ⁵anisah@atmaluhur.ac.id,

⁶syafurirawadi@atmaluhur.ac.id

Abstract

Jims honey is a product that is in great demand by all groups and is a current fashion trend, especially among teenagers and mothers. At Jims honey Pangkalpinang, which is located on Jl Darul Mahabbah, sales are currently very rapid, especially when sales are made online. At this time Jims Honey has become a lot of resellers in every district on Bangka Island. Sales in one month sold more than 1,000 products from a variety of Jims Honey products. The purpose of this research is to find out which products are the most selling to the lowest each month with the a priori algorithm. Where the stage of this research method is carried out first to obtain overall data for that month. The second stage is analysis using the apriodi algorithm, the third stage is applying the a priori algorithm system with data mining obtained at Jims Honey Pangkalpinang to find the best-selling to lowest data product results for that month.

Keywords: Apriori, Jims Honey, Data Mining, Section

Abstrak

Jims honey merupakan produk yang sangat diminati oleh semua kalangan dan menjadi trend fashion saat ini khusus di kalangan remaja dan ibu-ibu. Pada Jims honey Pangkalpinang yang bertempat di Jl Darul Mahabbah saat ini penjualan sangat pesat terutama saat penjualan dilakukan secara online. Pada Jims Honey saat ini telah banyak menjadi reseller-reseller ada di setiap kabupaten yang ada di Pulau Bangka. Penjualan dalam satu bulan produk terjual lebih dari 1.000 produk dari berbagai macam produk Jims Honey. Tujuan penelitian ini untuk mengetahui produk yang terlaris paling banyak hingga terendah setiap bulannya dengan algoritma apriori. Dimana tahap metode penelitian ini dilakukan pertama memperoleh data secara keseluruhan pada bulan tersebut. Tahap kedua analisa dengan algoritma apriodi, tahap ketiga menerapkan ke sistem algoritma apriori dengan data mining yang diperoleh pada Jims Honey Pangkalpinang untuk mencari hasil produk data terlaris hingga terendah pada bulan tersebut.

Kata kunci: Apriori, Jims Honey, Data Mining

1. Pendahuluan

Jims Honey merupakan salah satu brand Lokal di Indonesia dimana saat ini penjualan setiap bulannya lebih dari 1.000 produk yang ada di Jims Honey. Adapun produk yang dijual Jims Honey yaitu tas, jam tangan, termos, powerbank dan dompet. Masing-masing produk memiliki popularitas tinggi di masing-masing produk. Data mining merupakan proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (mechine learning) untuk menganalisis dan mengekstraksi pengetahuan (knowledge) secara otomatis [1]. Data Mining ialah proses penemuan hubungan antar pola kecenderungan

hubungan antar data dalam kumpulan yang besar di dalam penyimpanan dengan menggunakan pola tertentu sehingga ditemukan pengetahuan yang baru tidak terduga [2]. Dengan data mining ini digunakan untuk mencari produk Jims Honey pada Kota Pangkalpinang yang terlaris dibeli pada bulan tersebut. Sehingga data tersebut mempermudah owner untuk memesan produk Jims Honey. Dari data yang ada di kelola dengan algoritma Apriori. Apriori ialah Algoritma Apriori adalah suatu algoritma dasar yang diusulkan oleh Agrawal & Srikant pada tahun 1994 untuk penentuan frequent itemsets untuk aturan asosiasi boolean [3], termasuk jenis aturan asosiasi pada data mining [4].

2. Metodologi Penelitian

Dalam penelitian ini membahas tentang data mining dengan menggunakan algoritma Apriori.

2.1. Data Mining

Data mining atau knowledge discovery in database (KDD) adalah proses resourcing dan penggunaan data untuk menemukan pola atau hubungan dari sekumpulan data berukuran besar [5]. data mining yang mengeksekusi operasi data mining yang telah didefinisikan berdasarkan model analisis. Data mining merupakan proses analisis terhadap data dengan penekanan menemukan informasi yang tersembunyi pada sejumlah data besar yang disimpan ketika menjalankan bisnis perusahaan [6].

2.2. Algoritma Apriori

Algoritma apriori adalah algoritma yang paling terkenal untuk menemukan pola frekuensi tinggi. Apriori dibagi menjadi beberapa tahap disebut narasi atau pass [7].

- a. Pembentukan kandidat itemset, kandidat k- itemset dibentuk dari kombinasi (k-1)-itemset yang didapat dari iterasi sebelumnya. Satu cara dari algoritma apriori adalah adanya pemangkasan kandidat k-itemset yang subset-nya yang berisi k-1 item tidak termasuk dalam pola frekuensi tinggi dengan panjang k1.
- b. Penghitungan support dari tiap kandidat k- itemset. Support dari tiap kandidat k- itemset didapat dengan menscan database untuk menghitung jumlah transaksi yang memuat semua item didalam kandidat k-itemset tersebut. Ini adalah juga ciri dari algoritma apriori dimana diperlukan penghitungan dengan cara seluruh database sebanyak k-itemset terpanjang.

Tetapkan pola frekuensi tinggi. Pola frekuensi tinggi yang memuat k item atau k-itemset ditetapkan dari kandidat k-itemset yang supportnya lebih besar dari minimum support. Bila tidak didapat pola frekuensi tinggi baru maka seluruh proses dihentikan. Bila tidak, maka k ditambah satu dan kembali bagian 1.

3. Hasil dan Pembahasan

Data transaksi penjualan selama sebulan sebanyak 140 transaksi dan Jumlah produk yang diperjual belikan adalah 109 produk dari total item 1.120. ketika di hitung maka dapat hasil seperti berikut.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
0	ANITA	JULIE	TUMBLR	JHW10	JHW09	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	EVELINE	FANIE	CLARESTA	JHW10	JHW09	ANAS	GENIE	MONIC	TUMBLR	NaN								
2	JHW10	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	TUMBLR	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	JHW10	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Gambar 1. Data Transaksi

Dari data diatas dapat diketahui jika pembelian produk terbanyak dalam satu kali transaksi adalah 17 produk seperti data di bawah ini

```
[[ 'ANITA', 'JULIE', 'TUMBLR', 'JHW10', 'JHW09' ],
 [ 'EVELINE',
  'FANIE',
  'CLARESTA',
  'JHW10',
  'JHW09',
  'ANAS',
  'GENIE',
  'MONIC',
  'TUMBLR' ],
 [ 'JHW10' ],
 [ 'JHW10' ],
 [ 'TUMBLR' ]]
```

Adapun transformasi data untuk diproses yaitu seperti Gambar 2.

	2139	8120	8123	8138	8151	8416	9123	AGNES	AIRA	AIRIN	...	TUMBLR	VERONICA	VINKA	VIO	VIORA	VIVIAN	VONY	WENDY	WINNIE	XAVIER		
0	False	...	True	False	False	False	False	False	False	False	False	False	False										
1	False	...	True	False	False	False	False	False	False	False	False	False	False	False									
2	False	...	False	False	False	False	False	False	False	False	False	False	False	False									

3 rows x 109 columns

Gambar 2. Transformasi Data

Dari Gambar 2 maka minimum support = $3x \text{ transaksi} / 140(\text{jumlah transaksi}) * 100 = 2.1\% (0.021)$

Dengan minimum confidence = 75% (0.75)

Untuk mencari nilai support sebuah item / produk

$\text{support}(A) = \text{Jumlah Transaksi yang Ada produk A} / \text{Total transaksi}$

$\text{support}(A \rightarrow B) = \text{Jumlah Transaksi yang ada produk (A, B)} / \text{Total transaksi}$

a) Analisis Pola frekuensi Tinggi

Tahap ini mencari pola kombinasi item yang memenuhi syarat minimum dari nilai *support* dalam database. Nilai *support* sebuah item diperoleh dengan rumus berikut:

$$\text{Support}(A) = \frac{\text{Jumlah transaksi mengandung A}}{\text{Total transaksi}} \quad (1)$$

Sedangkan nilai *support* dari 2 item diperoleh dari :

$$\text{Support}(A, B) = P(A \cap B)$$

$$\text{Support}(A, B) = \frac{\epsilon \text{Jumlah transaksi mengandung A dan B}}{\epsilon \text{Total transaksi}} \quad (2)$$

b) Pembentukan aturan Asosiasi

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence iaturan asosiatif* $A \rightarrow B$

$$\text{Confidence} = P\left(\frac{B}{A}\right) = \frac{\epsilon \text{ Transaksi mengandung A dan B}}{\epsilon \text{ Transaksi mengandung A}} \quad (3)$$

Minimum support 0,021 (2.1%) dari transaksi

	support	itemsets	length
0	0.150000	(8123)	1
1	0.028571	(8138)	1
2	0.050000	(8416)	1
3	0.035714	(AGNES)	1
4	0.028571	(AIRA)	1
...
138	0.021429	(GENIE, MONIC, EILEEN)	3
139	0.021429	(MISHA, EILEEN, JHW09)	3
140	0.021429	(TUMBLR, EILEEN, MISHA)	3
141	0.021429	(TUMBLR, EVELINE, JHW09)	3
142	0.028571	(TUMBLR, JHW10, JHW09)	3

143 rows x 3 columns

Gambar 3. Data Support

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	antecedent_len
0	(ALICIA)	(SYLVIA)	0.035714	0.078571	0.028571	0.80	10.181818	0.025765	4.607143	1
1	(LAYLA)	(AMBER)	0.028571	0.100000	0.021429	0.75	7.500000	0.018571	3.600000	1
2	(ANITA)	(TUMBLR)	0.028571	0.121429	0.021429	0.75	6.176471	0.017959	3.514286	1
3	(MOANA)	(CHLOE)	0.028571	0.114286	0.021429	0.75	6.562500	0.018163	3.542857	1
4	(JOICE)	(EVELINE)	0.028571	0.078571	0.021429	0.75	9.545455	0.019184	3.685714	1
5	(FANIE)	(MONIC)	0.021429	0.064286	0.021429	1.00	15.555556	0.020051	inf	1
6	(LAYLA)	(GENIE)	0.028571	0.128571	0.021429	0.75	5.833333	0.017755	3.485714	1
7	(SYLVIA, OLIVIA)	(8123)	0.028571	0.150000	0.021429	0.75	5.000000	0.017143	3.400000	2
8	(8123, OLIVIA)	(SYLVIA)	0.021429	0.078571	0.021429	1.00	12.727273	0.019745	inf	2
9	(ANNA, EVELINE)	(CHLOE)	0.021429	0.114286	0.021429	1.00	8.750000	0.018980	inf	2
10	(CHLOE, JHW09)	(EVELINE)	0.021429	0.078571	0.021429	1.00	12.727273	0.019745	inf	2
11	(JOAN, EVELINE)	(CHLOE)	0.021429	0.114286	0.021429	1.00	8.750000	0.018980	inf	2
12	(JOAN, CHLOE)	(EVELINE)	0.028571	0.078571	0.021429	0.75	9.545455	0.019184	3.685714	2
13	(JULIE, CHLOE)	(EVELINE)	0.021429	0.078571	0.021429	1.00	12.727273	0.019745	inf	2
14	(JULIE, EVELINE)	(CHLOE)	0.021429	0.114286	0.021429	1.00	8.750000	0.018980	inf	2
15	(GENIE, MONIC)	(EILEEN)	0.028571	0.121429	0.021429	0.75	6.176471	0.017959	3.514286	2
16	(GENIE, EILEEN)	(MONIC)	0.028571	0.064286	0.021429	0.75	11.666667	0.019592	3.742857	2
17	(MONIC, EILEEN)	(GENIE)	0.028571	0.128571	0.021429	0.75	5.833333	0.017755	3.485714	2
18	(JHW09, MISHA)	(EILEEN)	0.021429	0.121429	0.021429	1.00	8.235294	0.018827	inf	2
19	(EILEEN, JHW09)	(MISHA)	0.021429	0.128571	0.021429	1.00	7.777778	0.018673	inf	2
20	(TUMBLR, MISHA)	(EILEEN)	0.028571	0.121429	0.021429	0.75	6.176471	0.017959	3.514286	2
21	(TUMBLR, EVELINE)	(JHW09)	0.021429	0.092857	0.021429	1.00	10.769231	0.019439	inf	2
22	(TUMBLR, JHW10)	(JHW09)	0.035714	0.092857	0.028571	0.80	8.615385	0.025255	4.535714	2
23	(JHW10, JHW09)	(TUMBLR)	0.028571	0.121429	0.028571	1.00	8.235294	0.025102	inf	2

Gambar 4. Data Hasil

Dari Gambar 4 dapat dibaca penjelasannya sebagai berikut yaitu jika beli alicia maka akan beli sylvia, jika beli layla maka akan beli amber, jika beli anita akan beli tumblr, jika beli sylvia dan olivia maka akan beli 8123 dan seterusnya.

4. Kesimpulan

Data mining dengan algoritma apriori sangat berguna penjualan JimsHoney yang paling diminati oleh konsumen, dimana dapat dijadikan informasi dalam pengambilan keputusan untuk stok produk bulan berikutnya. Berdasarkan hasil penelitian ini yang paling diminati yaitu fanie.

Daftar Pustaka

- [1] V. N. Budiyasari, "Implementasi Data Mining Pada Penjualan kacamata Dengan Menggunakan Algoritma Apriori," *Indones. J. Comput. Inf. Technol*, vol. 2, no. 2, pp. 31–39, 2017.
- [2] M. Khanza and R. Teyib, "Implementasi Algoritma Apriori Dalam Penentuan Pemesanan Barang Untuk Transaksi Penjualan Handphone," *J. Sci. Appl. Informatics*, vol. 4, no. 2, pp. 221–235, 2021.
- [3] F. Rahmawati and N. Merlina, "Metode Data Mining Terhadap Data Penjualan Sparepart Mesin Fotocopy Menggunakan Algoritma Apriori," *PIKSEL Penelit. Ilmu Komput. Sist. Embed. Log*, vol. 6, no. 1, pp. 9–20, 2018.
- [4] E. D. Sikumbang, "Penerapan Data Mining Penjualan Sepatu Menggunakan Metode Algoritma Apriori," *J. Tek. Komput. AMIK BSI*, vol. 8, no. 1, pp. 84–95, 2018.
- [5] P. N. Harahap and S. Sulindawaty, "Implementasi Data Mining Dalam Memprediksi Transaksi Penjualan Menggunakan Algoritma Apriori (Studi Kasus PT.Arma Anugerah Abadi Cabang Sei Rampah)," *Matics*, vol. 11, no. 2, p. 46, 2020, doi: 10.18860/mat.v11i2.7821.
- [6] F. Rahmawati and N. Merlina, "Metode Data Mining Terhadap Data Penjualan Sparepart Mesin Fotocopy Menggunakan Algoritma Apriori," *PIKSEL Penelit. Ilmu Komput. Sist. Embed. Log.*, vol. 6, no. 1, pp. 9–20, 2018, doi: 10.33558/piksel.v6i1.1390.
- [7] F. A. K. Wardani and T. Kristiana, "Implementasi Data Mining Penjualan Produk Kosmetik Pada PT. Natural Nusantara Menggunakan Algoritma Apriori," *Paradig. - J. Komput. dan Inform.*, vol. 22, no. 1, pp. 85–90, 2020, doi: 10.31294/p.v22i1.6520.